



Jena Research Papers in Business and Economics

Zur Berücksichtigung von Unsicherheit bei der Präferenzmessung

Roland Helm und Armin Scholl

24/2007

Jenaer Schriften zur Wirtschaftswissenschaft

Working and Discussion Paper Series
School of Economics and Business Administration
Friedrich-Schiller-University Jena

ISSN 1864-3108

Publisher:

Wirtschaftswissenschaftliche Fakultät
Friedrich-Schiller-Universität Jena
Carl-Zeiß-Str. 3, D-07743 Jena
www.jbe.uni-jena.de

Editor:

Prof. Dr. Hans-Walter Lorenz
h.w.lorenz@wiwi.uni-jena.de
Prof. Dr. Armin Scholl
armin.scholl@wiwi.uni-jena.de

www.jbe.uni-jena.de

Zur Berücksichtigung von Unsicherheit bei der Präferenzmessung

Anmerkungen zum Beitrag „Präferenzmessung bei unsicheren Produkteigenschaften: Risikoberücksichtigung bei Ergebnissen aus Conjoint-Analysen“

Roland Helm^a und Armin Scholl^b

Zusammenfassung

Der Beitrag beschäftigt sich mit der Problematik, dass im Rahmen von Präferenzmessungen Ausprägungen einzelner Merkmale unsicher sein können. Die möglichen negativen Auswirkungen dieser Unsicherheit und Ansätze zu ihrer Beseitigung im Rahmen von Conjoint-Analysen werden in einem Beitrag von Nitschke und Völckner thematisiert. Die dort gemachten Vorschläge und Schlussfolgerungen werden hier einer kritischen Prüfung unterzogen. Dabei stellt sich heraus, dass die beschriebenen Effekte zwar bedeutsam sind, die vorgeschlagene Methodik jedoch nicht zu deren Berücksichtigung geeignet ist. Korrekturen und Alternativvorschläge werden unterbreitet.

^a Univ.-Prof. Dr. Roland Helm, Inhaber des Lehrstuhls für ABWL / Absatzwirtschaft, Marketing und Handel, Friedrich-Schiller-Universität Jena, Carl Zeiß-Straße 3, 07743 Jena, Tel. 03641/943110, eMail: roland.helm@wiwi.uni-jena.de

^b Univ.-Prof. Dr. Armin Scholl, Inhaber des Lehrstuhls für ABWL / Betriebswirtschaftliche Entscheidungsanalyse, Friedrich-Schiller-Universität Jena, Carl Zeiß-Str. 3, 07743 Jena, Tel. 03641/943170, eMail: armin.scholl@wiwi.uni-jena.de

Zur Berücksichtigung von Unsicherheit bei der Präferenzmessung

Anmerkungen zum Beitrag „Präferenzmessung bei unsicheren Produkteigenschaften: Risikoberücksichtigung bei Ergebnissen aus Conjoint-Analysen“

1. Disposition

Im Beitrag „Präferenzmessung bei unsicheren Produkteigenschaften: Risikoberücksichtigung bei Ergebnissen aus Conjoint-Analysen“ von *Thomas Nitschke* und *Franziska Völckner* in der Ausgabe der *Zeitschrift für betriebswirtschaftliche Forschung* vom September 2006 stellen die Autoren einen Effekt in den Mittelpunkt der Betrachtungen, der bei der Präferenzmessung grundsätzlich von Bedeutung ist.¹

Konkret geht es um die Messung der Präferenzstrukturen potenzieller Nachfrager durch die Conjoint-Analyse unter Berücksichtigung der Unsicherheit, die aus Konsumentensicht durch das Vorliegen nur allgemeiner Angaben über die Merkmalsausprägungen verursacht wird. Gleichwohl könnte auch in der Realität die Bewertung der Nutzen konkreter Alternativen am Markt unsicher sein. Die Autoren argumentieren, dass die Konsumenten interindividuell verschiedene Vermutungen über die mittleren (aber nicht nur über diese, sondern gegebenenfalls über alle) Merkmalsausprägungen sowie die „Streuung der Eintrittswahrscheinlichkeiten“ haben können (S. 744). Sie meinen damit offensichtlich die interindividuell unterschiedlich eingeschätzten Ausprägungen und Ausprägungsbandbreiten realer Alternativen. Die Conjoint-Analyse spiegele die „empfundene Unsicherheit (...) jedoch nicht ausreichend wider“, Produktkonzepte würden als sicher präsentiert (S. 745). Eigenschaften seien in der Realität jedoch teilweise unsicher, so dass die Präferenzmessung eben diesen realiter gegebenen Sachverhalt auch abbilden müsse. Dies sei bisher nicht gegeben, einzelne nicht näher beschriebene Ansätze ausgenommen (S. 746, Fußnote 10). Ein entsprechender neuer Ansatz, diese Unsicherheit im Nachhinein durch Modifikation der ermittelten Teilnutzenwerte zu berücksichtigen, solle vorgestellt werden (S. 746).

¹ Vgl. *Nitschke/Völckner* (2006).

Die Ausführungen bedürfen einiger Klarstellungen und Berichtigungen. Zum einen ist außerdem an verschiedenen Stellen die Verständlichkeit eingeschränkt, wodurch sich einige offene Fragen ergeben. Zum anderen enthält der Artikel formale Mängel.

Folgende Punkte erscheinen am wichtigsten:

1. Teilweise setzen die Autoren miteinander unvereinbare Begrifflichkeiten synonym, andere Begriffe werden nicht klar definiert. Außerdem werden mathematische Zusammenhänge an verschiedenen Stellen formal unsauber beschrieben. Mit diesen Kritikpunkten setzen wir uns in Abschnitt 2 auseinander.
2. Die Autoren stellen eine Methode zur Analyse und nachträglichen Behebung des festgestellten Problems vor. Aufgrund von Ungenauigkeiten und Auslassungen bei deren Beschreibung ist unklar, ob diese Methode fehlerhaft oder das Problem gar nicht existent ist. In Abschnitt 3 wird auf diese beiden Fälle näher eingegangen, während in Abschnitt 4 einfache und nahe liegende Alternativvorschläge zur tatsächlichen Problembehandlung in ihren Grundzügen dargelegt werden.
3. Alternative Möglichkeiten der Problembehandlung bzw. grundsätzliche Prämissen der Präferenzmessung werden nicht angesprochen bzw. die Frage der Notwendigkeit von Korrekturmaßnahmen wird ganz generell nicht diskutiert. Zu diesen Fragen nehmen wir in Abschnitt 5 in aller Kürze Stellung.
4. Die durchgeführte Studie leidet über die Fragwürdigkeit der Methodik hinaus unter weiteren Mängeln, die in Abschnitt 6 diskutiert werden.

Der Aussagegehalt des Beitrags ist daher in erheblichem Ausmaß eingeschränkt bzw. zu hinterfragen. Dies soll im Folgenden weiter vertieft werden.

2. Zur Terminologie und zu formalen Fehlern

Die Autoren sprechen von einer „Streuung der Eintrittswahrscheinlichkeiten“ (S. 744, 747) und gar von einem Mittelwert derselben (S. 747). Dies deutet auf eine zweistufige Unsicherheit hin, nicht nur die Ausprägungen sind unsicher, sondern auch die Wahrscheinlichkeiten ihres Auftretens. Vermutlich wird jedoch schlicht auszudrücken versucht, dass die Ausprägungen selbst stochastisch sind, also einer Wahrscheinlichkeitsverteilung unterliegen. Eine weitere begriffliche Ungenauigkeit besteht bei der Verwendung der Begriffe Erwartungswert und Erwartungsnutzen (S. 748 und S. 752).

Es wird im gesamten Artikel von der Unsicherheit der Eigenschaftsausprägungen gesprochen. Dabei ist offensichtlich nicht gemeint, dass die Ausprägungsstufen als solche unsicher oder ungenau beschrieben sind (Was ist zum Beispiel eine „mittlere“ Video-Qualität?), sondern dass ungewiss ist, welche der bekannten Ausprägungsstufen einer

Eigenschaft bei einem konkreten Angebot tatsächlich auftritt (Wie gut ist zum Beispiel die Qualität eines konkreten Downloadangebotes wirklich, das als „mittel“ eingestuft wird?). Nun wird kritisiert (S. 745), dass die Conjoint-Analyse diese letztere Form der Unsicherheit außer acht lässt, indem sie allgemein wie folgt vorgeht: Den Probanden wird gezeigt, was unter den verschiedenen Ausprägungsstufen zu verstehen ist, und es werden ihnen Stimuli zur Bewertung vorgelegt, die aus diesen dann sicheren Ausprägungen der verschiedenen Merkmale zusammengesetzt sind. Es ist somit aus Sicht der Probanden überhaupt keine Unsicherheit mehr in den Bewertungsaufgaben enthalten (konkrete Ausprägungen mit bekannter Interpretation). Diese Vorgehensweise entspricht dem Hauptzweck der Conjoint-Analyse, da man ja genau die verschiedenen Ausprägungen von *bekannt* Produkteigenschaften nutzenmäßig bewerten möchte.² Insofern zielt die Kritik an der Conjoint-Analyse nicht auf die Methode selbst, sondern allenfalls auf deren allgemein akzeptierte Annahmen, die Art der Anwendung bei Nichtvorliegen dieser Anwendungsvoraussetzungen und/oder auf ganz spezielle Auswertungsmöglichkeiten mit verfügbarer Software.

Bevor die vorgeschlagene Methode diskutiert wird, soll eine Unterscheidung verschiedener Arten von Unsicherheit zum besseren Verständnis beitragen; eine derartige Klassifikation wäre bei *Nitschke/Völckner* (2006) aus unserer Sicht ebenfalls notwendig gewesen:

- *Bedeutungsunsicherheit (Unschärfe, Fuzziness)*: In wie weit ist dem Probanden überhaupt klar, was eine bestimmte Ausprägungsstufe eines Merkmals bedeutet? So ist zum Beispiel zu klären, was der Proband unter einer mittleren Videoqualität zu verstehen hat. Dies wird in der Conjoint-Analyse, wie korrekt beschrieben, durch Präsentation der Eigenschaftsausprägungen auf geeignete Weise vorgenommen. Dadurch verbleibt im Rahmen einer korrekt ausgeführten Conjoint-Analyse keine Bedeutungsunsicherheit. Dies haben die Autoren in ihrer Studie entsprechend umgesetzt (S. 757).
- *Ausprägungsunsicherheit (Ergebnisunsicherheit)*: Hierunter verstehen wir den klassischen, wahrscheinlichkeitstheoretischen Fall der Unsicherheit. Man weiß, welche Szenarien bestehen, d.h. hier, die möglichen Ausprägungen sind bedeutungssicher. Man weiß allerdings nicht, welches Szenario eintritt, allenfalls sind Wahrscheinlichkeiten bekannt (entscheidungstheoretischer Fall „Risiko“). Im betrachteten Kontext bedeutet dies zum Beispiel, dass der Proband nicht weiß, welche der bedeutungssicheren Ausprägungsstufen er tatsächlich bekommt, wenn die Ausprägungsstufe „mittel“ für die Videoqualität angegeben ist. Er glaubt also dem Angebot nur begrenzt und kann allenfalls Wahrscheinlichkeiten für die tatsächlichen Ausprägungen prognostizieren. Mit derartigen Problemen

² Vgl. *Backhaus/Erichson/Plinke/Weiber* (2003), S. 544.

beschäftigt sich die Entscheidungstheorie mit Hilfe geeigneter Maße der zugrundeliegenden Wahrscheinlichkeitsverteilung. Dies war offensichtlich der Ansatzpunkt für *Nitschke* und *Völckner* mit dem Ziel, entscheidungstheoretische Ansätze in die Conjoint-Analyse zu integrieren.

- *Nutzenunsicherheit (Bewertungsunsicherheit)*: Hierbei besteht Unsicherheit über die eigenen Präferenzvorstellungen. So ist es etwa nicht möglich, den mit der Ausprägungsstufe „mittlere Qualität“ verbundenen Nutzen eindeutig zu benennen. Verlangt eine Methode, wie auch die Conjoint-Analyse, nun, dass ein genauer Nutzenwert benannt wird, führt dies – ebenso wie eine Ausprägungsunsicherheit – zu falschen bzw. durch Einengung auf einen festen Wert unnötig harten Präferenzaussagen. Um diesem Problem zu begegnen, sollte man den Auskunftspersonen erlauben, ihre Unsicherheit auszudrücken, zum Beispiel durch die Angabe von Nutzenintervallen oder weichen Präferenzrelationen wie zum Beispiel bei Outranking-Verfahren.³

Die Unterscheidung und Diskussion verschiedener Arten von Unsicherheit, die Auswirkungen auf eine Präferenzmessung haben können, zeigen zum einen, dass es wohl die Ausprägungsunsicherheit ist, auf die sich *Nitschke* und *Völckner* beziehen (auch wenn sich diese Erkenntnis den Autoren dieser Anmerkungen erst nach mehrmaligem Lesen des Artikels erschlossen hat). Zum anderen wird aber auch klar, dass weitere Formen der Unsicherheit für falsche Präferenzmessungen sorgen können. So ist es insbesondere im Rahmen der durchgeführten Studie (S. 754ff.) kaum möglich, Abweichungen allein an der Ausprägungsunsicherheit festzumachen.

Die formale Darstellung der durch die (Ausprägungs-) Unsicherheit verursachten Fehlmessungen (unterteilt in den ersten und zweiten Unsicherheits-Bias) im Rahmen der Conjoint-Analyse ist an verschiedenen Stellen formal unsauber und daher schwer verständlich (S. 748-753). Einige der verwendeten Parameter werden nicht oder nicht in der verwendeten Weise definiert, mathematische Zusammenhänge sind nicht immer korrekt dargestellt. Da wir jedoch die gewählte Vorgehensweise insgesamt für fragwürdig halten, sollen derartige formale Details nicht weiter vertieft werden.

3. Zur nachträglichen Berücksichtigung von Unsicherheit

Da in *Nitschke/Völckner* (2006) entscheidende Aussagen nicht hinreichend deutlich sind und zwei Problemkreise einander überlagern, werden im Folgenden zwei Argumentationslinien verfolgt.

³ Vgl. *Klein/Scholl* (2004), Kap. 7.3.1.2.

Zum einen wird unterstellt, dass die Unsicherheit über Eigenschaftsausprägungen im Rahmen der Conjoint-Analyse eine Rolle spielt (insbesondere aufgrund der Anlage der empirischen Studie), so dass der Ansatz von *Nitschke* und *Völckner* die durch diese Unsicherheit verursachten Verzerrungen bzw. Fehler nachträglich korrigieren müsste.

Zum anderen diskutieren wir den Fall, dass Unsicherheit im Rahmen der Conjoint-Analyse selbst keine Rolle spielt, indem dort Bedeutungs- und Ausprägungssicherheit herbeigeführt werden. In diesem Fall korrigiert die Methode lediglich solche Verzerrungen, die durch einen überflüssigen Verfahrensschritt verursacht werden, aber eigentlich gar nicht bestehen.

Fall 1: Ausprägungsunsicherheit im Rahmen der Conjoint-Analyse

Eine genaue und eindeutige Beschreibung der Attribute und deren Ausprägungen⁴ ist Grundvoraussetzung jeder Präferenzmessung.⁵ Es muss sichergestellt sein, dass jede Auskunftsperson (jeder Proband) unter den verwendeten Begriffen das Gleiche versteht (Bedeutungssicherheit) und tatsächlich die vorgegebenen Ausprägungsstufen erwartet (Ausprägungssicherheit). Nur dann ist ein Vergleich über die Auskunftspersonen und eine Aggregation möglich. Insofern erfüllt die von *Nitschke* und *Völckner* untersuchte Problemstellung diese grundlegenden Anforderungen nicht, da aufgrund der unterstellten Ausprägungsunsicherheit kaum einer der Probanden gerade die Ausprägung bewerten wird, die man ihm zur Bewertung vorlegt (auch wenn Bedeutungssicherheit durch Zeigen von Beispielen hergestellt werden sollte). In einer solchen Situation darf die Conjoint-Analyse gar nicht angewendet werden und muss zwingend fehlerhafte Ergebnisse liefern, die sich im Nachhinein auch nicht mehr korrigieren lassen, wie dies die Autoren offenkundig versuchen. Dies ist eine relativ triviale Erkenntnis, die im Folgenden genauer ausgeführt wird:

Wenn man einem Probanden zum Beispiel einen Stimulus „Tauschbörse mit mittlerer Qualität“ vorlegt und er befürchtet, nicht die - zur Gewährleistung der Bedeutungssicherheit vorgeführte - Qualität „mittel“ zu bekommen, so bewertet er tatsächlich den Stimulus „Tauschbörse mit individuell erwarteter Qualität“, die zum Beispiel zwischen der objektiv bekannten „mittleren“ und „hohen“ Qualität liegt. Werden nun derartige Stimuli im Rahmen der Conjoint-Analyse direkt bewertet, in eine Rangfolge gebracht oder paarweise verglichen, haben die gewonnenen Erkenntnisse keinerlei Bedeutung für die Bewertung der vorgelegten Stimuli. Stattdessen bewerten sie die vom jeweiligen Probanden unterstellten Stimuli. Kennt man diese nicht, wie hier unterstellt, kann man keine sinnvollen Nutzenberechnungen vornehmen.

⁴ Vgl. *Shocker/Srinivasan* (1979), S. 14.

⁵ Vgl. *Keeney/Gregory* (2005), S. 3f.; *Payne/Bettman/Schkade* (1999), S. 245.

Tut man es dennoch, entsteht der „erste Unsicherheitsbias“ (S. 748), wie korrekt von den Autoren erkannt. Ein solcher ergibt sich aber ebenfalls, wenn man dem Probanden verdeckt einen als Birne bezeichneten Apfel zu essen gibt und an seinen Aussagen zum Geschmack eine Bewertung für Birnen vornimmt. Bei einer solchen Vorgehensweise kann sich bekanntlich nichts Sinnvolles ergeben. Vor allem ist es nicht möglich, aus dem Geschmack der Äpfel durch eine, wie auch immer geartete, nachträgliche Transformation Aussagen über den Geschmack von Birnen herzuleiten. Aber genau dies versuchen die Autoren durch die Berechnungen auf Seite 748ff. Was macht eine Conjoint-Analyse, die diese Problematik ignoriert, also falsch? Es werden Nutzeneinschätzungen schlicht mit den falschen Stimuli rechnerisch verbunden. Letztlich entspricht dies genau dem Problem, das bei einer nicht beseitigten Bedeutungsunsicherheit vorliegen würde. In einem solchen Fall würde man die Conjoint-Analyse aufgrund nicht bestehender Anwendungsvoraussetzungen ebenfalls nicht ohne weitere vorbereitende Schritte einsetzen können.

Als Folge dieser falschen Messung haben die Ergebnisse der Conjoint-Analyse keine sinnvolle Aussagekraft, da der Teilnutzenwert $u_k(r_{\tilde{h}})$ einer Ausprägungsstufe $r_{\tilde{h}}$ für den Konsumenten k (S. 750) nicht den so bezeichneten Teilnutzenwert, sondern eine nicht mehr definierbare „Mischung“ von Teilnutzenwerten ganz verschiedener Ausprägungsstufen darstellt (unter $r_{\tilde{h}}$ hat sich ja jeder Proband etwas anderes vorgestellt). Verwendet man diese Werte nun als Basis zur Berechnung von Nutzenabweichungen, wie in den Gleichungen (5) und (7) auf Seite 750 und 751 geschehen, ergeben sich keine sinnvollen Erkenntnisse. Stattdessen müsste man tatsächlich korrekte Teilnutzenwerte bereits kennen (dazu wird in Abschnitt 4 ein Vorschlag gemacht), wodurch man den tatsächlichen Fehler einer unsinnig angewendeten Conjoint-Analyse bestimmen und diesen Fehler vermeiden könnte. Die angegebenen Formeln zur Berechnung der Abweichungen sind formal (bis auf mathematische Unsauberkeiten an einzelnen Stellen) also durchaus korrekt, sie können nur nicht im Rahmen der Conjoint-Analyse in der beschriebenen Weise zur Korrektur der Teilnutzenwerte verwendet werden.

Der „zweite Unsicherheitsbias“ (S. 748) beschreibt die für Ausprägungsunsicherheit zusätzlich auftretende Problematik, dass der Proband bei Risikoscheu befürchtet, dass er nicht die erwartete Qualität (im Sinne des Erwartungswertes der individuellen Zufallsvariable „Qualität“), sondern eventuell noch etwas Schlechteres bekommt (zum Beispiel im Sinne eines Value-at-Risk). Umgekehrt ist es beim Risikofreudigen, wie auch Untersuchungen von *Gierl* (2006) über den Präferenzeinfluss von ambigen Ausprägungen belegen (leider wird dieser bedeutsame Unterschied nicht bereits auf S. 746, sondern erst auf S. 748 erläutert). Die aus der Risikoeinstellung resultierenden Abweichungen zwischen gemessenen und tatsächlichen Präferenzwerten sollen ebenfalls nachträglich herausgerechnet werden (S. 750ff.). Auch dies ist aufgrund der oben beschriebenen Fehlmessungen prinzipiell nicht (mehr) möglich. Außerdem erscheinen die willkürliche

Wahl der logarithmischen Risikonutzenfunktion und der damit verbundene Ansatz auf Seite 752f. mehr als fragwürdig.

Zusammenfassend muss gesagt werden, dass eine nachträgliche Korrektur der Teilnutzenwerte und damit der Gesamtnutzen zu bewertender Alternativen grundsätzlich nicht möglich ist, wenn zur Schätzung der Teilnutzenwerte auf eine für alle Probanden gleiche Ausprägungsstufe \bar{h} zurückgegriffen wird. Es ist stattdessen notwendig, die betrachteten Fehlerquellen im vorhinein zu berücksichtigen bzw. für jeden Probanden eine an seinen konkreten Ausprägungsstufen orientierte Schätzung vorzunehmen. Ein entsprechender Vorschlag wird in Abschnitt 4 unterbreitet.

Die bisherigen Ausführungen der ersten Argumentationslinie beruhen auf der Annahme, dass die Probanden im Rahmen der Conjoint-Analyse die Ausprägungsunsicherheit explizit empfinden und unter dieser Bedingung Präferenzeinschätzungen abgeben. Diese Vermutung erscheint aufgrund der Gestaltung der Studie mehr als berechtigt. In einem ersten Schritt werden die Probanden in einer schriftlichen Befragung gerade mit den Themen Online-Distribution von Videos (inkl. der Tauschbörsen) und Risiko konfrontiert (S. 755 unten). Außerdem wurden nur Probanden gewählt, die Erfahrungen mit Tauschbörsen, also auch der dort unsicheren Qualität hatten (S. 755 Mitte). Dann präsentiert man ihnen im Rahmen der Conjoint-Analyse Stimuli, die unter anderem auch den Fall der Tauschbörse (repräsentiert durch einen Preis von 0 Euro; vgl. Tabelle 1 und Tabelle A-1) beinhalten. Auch wenn man vorher durch Präsentieren der Qualitätsstufen „hoch“, „mittel“ und „niedrig“ Bedeutungssicherheit erfolgreich hergestellt haben sollte (S. 755 unten), werden die Probanden, gerade aufgrund der Sensibilisierung durch die Vorbefragung, ihre Unsicherheit im Fall der Tauschbörsen bei ihren Präferenzeinschätzungen einfließen lassen.

So wie es aus Sicht der Verfasser dieser Anmerkungen aus *Nitschke/Völckner* (2006) herausgelesen werden kann, wird die Conjoint-Analyse (zumindest in der Studie) nicht unter der notwendigen Anwendungsvoraussetzung der Bedeutungs- und Ausprägungssicherheit durchgeführt. Die daraus entstehenden fehlerhaften Teilnutzen können mit der vorgeschlagenen Methode grundsätzlich nicht mehr korrigiert werden.

Fall 2: Ausprägungssicherheit im Rahmen der Conjoint-Analyse

Um einen zweiten Problemkreis zu diskutieren und der gebotenen Fairness gegenüber den Autoren Rechnung zu tragen, nehmen wir nun an, dass die Aussage zur (Bedeutungs-) Sicherheit der Ausprägungsstufen der Qualität auf Seite 755 („... dass alle Probanden genau einschätzen können, was sie bei einer „hohen“, „mittleren“ und niedrigen Bildqualität ... erhalten“) auch die Ausprägungssicherheit mit einschließen soll. Außerdem unterstellen wir entgegen obiger Argumentation, dass die Studie so angelegt wäre, dass die Conjoint-Analyse selbst unter vollständigem Ausschluss von Unsicherheit statt-

finden könnte. In diesem Fall erhält man für alle relevanten Ausprägungsstufen tatsächliche Teilnutzenwerte auf individueller Ebene, die sich unter bestimmten Bedingungen sinnvoll aggregieren lassen, wie dies im Rahmen von Conjoint-Analysen regelmäßig geschieht. Außerdem lässt sich jedes konkrete (Produkt-) Angebot durch Aufaddieren der Teilnutzen der dort vorhandenen Ausprägungen zu einem Gesamtnutzen bewerten und das beste Angebot ermitteln.

Ist die Ausprägung bei (mindestens) einem Merkmal unsicher, so ist ein sinnvoller Ansatz sicherlich der, für dieses Merkmal eine erwartete Ausprägungsstufe (im Sinne des Erwartungswertes der Verteilung der möglichen Ausprägungsstufen) einzusetzen. Dies kann – wie vorgeschlagen - durch Erfragen dieser Verteilung und Berechnen des Erwartungswertes gemäß Formel (6) auf Seite 750 vorgenommen werden (wenn man davon absieht, dass ein Erwartungswert ordinaler Größen nicht definiert ist; vgl. dazu Abschnitt 6).

Nun argumentieren *Nitschke* und *Völckner*, dass man eine einheitliche Ausprägungsstufe \bar{h} für alle Probanden einsetzen müsse, um den Kundennutzen eines Angebots mit ausprägungsunsicherer Ausprägung zu bestimmen (S. 749 und 758). Ganz offensichtlich ergibt sich dadurch genau das weiter oben beschriebene Problem des Vergleichs von Äpfeln und Birnen. Entsprechend sind die berechneten Gesamtnutzen der zu bewertenden Angebote bei allen Probanden falsch, die nicht zufällig genau diese Durchschnittsausprägungen erwarten. Die gesamten Berechnungen der Seiten 748 bis 751 dienen lediglich dem Zweck, diesen Fehler (erster Unsicherheits-Bias) zu berechnen. Nun ist dies geradezu tautologischer Natur, wie in der Argumentation zu Fall 1 nachgewiesen: Setzt man in eine Funktion das falsche Argument ein, erhält man in aller Regel einen falschen Funktionswert. Dies bedarf keiner aufwändigen Beweisführung.

Endgültig fragwürdig wird die gesamte formale Analyse jedoch, wenn man erkennt, dass dieser vorgeschlagene Schritt der Verwendung einer einheitlichen Ausprägungsstufe für alle Probanden völlig überflüssig ist. Wieso setzt man ein falsches Argument ein, wenn man das richtige kennt? Es wird behauptet, dass dies für das „klassische Vorgehen“ der Conjoint-Analyse notwendig sei (S. 747, 749 und 758). Dies wird jedoch nicht stichhaltig belegt. Nach unserer Auffassung kann davon auch keine Rede sein. Liegen individuelle Teilnutzen und die zu bewertenden (erwarteten) Ausprägungen vor, so gibt es keinen Grund, zunächst eine Generalisierung und Aggregation vorzunehmen und erst dann den dann falschen Gesamtnutzen zu berechnen. Es mag vielleicht mit verfügbarer Software keine einfache Möglichkeit geben, Gesamtnutzen individualisiert zu bestimmen (in Fußnote 11 angedeutet), aber dies kann kein Grund sein, einen solchen methodischen Fehler zu begehen. Die Berechnung der individuell korrekten Nutzenwerte lässt sich mit jeder Tabellenkalkulationssoftware auf einfachste Weise bewerkstelligen.

Natürlich folgt daraus, dass man bei individuell unterschiedlichen Ausprägungsstufen eine intuitive Form der Aggregation über alle Probanden nicht vornehmen darf. Aber diese ist auch keinesfalls notwendig. Stattdessen sollte man auf individueller Basis oder durch Gruppierung ähnlicher Probanden Erkenntnisse über Zahlungsbereitschaften und Vorhersagen über das Marktverhalten ableiten.

Zusammenfassend sei gesagt, dass die Autoren zu Recht anmerken, dass eine Aggregation der Teilnutzen bei Unsicherheit problematisch bzw. falsch ist; sie ziehen daraus jedoch die falschen Schlüsse bzw. sie weisen lediglich nach, dass ein offensichtlich falsches Vorgehen tatsächlich falsch ist. Dafür ist jedoch kein Artikel in einer anerkannten wissenschaftlichen Zeitschrift notwendig, es sei denn dieser Fehler werde regelmäßig in realen Studien tatsächlich begangen. Dies müsste jedoch anhand von Quellen besser belegt werden.

4. Alternativvorschlag zur Berücksichtigung der Unsicherheit

Unter der Annahme, dass eine explizite Berücksichtigung der Unsicherheit bei der Präferenzmessung tatsächlich notwendig ist (vgl. Abschnitt 5), kann man zum Beispiel entlang folgender Überlegungen vorgehen. Es soll dabei lediglich gezeigt werden, dass eine Unsicherheitsbehandlung prinzipiell auf formal korrekte Weise möglich ist. Für die konkrete Ausgestaltung und Analyse eines solchen Ansatzes wären weitere Forschungen nötig.

Um die Ausprägungsunsicherheit bereits während der Präferenzmessung und nicht nachträglich berücksichtigen zu können, ist eine Version der Conjoint-Analyse zu verwenden, die individuelle Auswertungen erlaubt. Dazu eignet sich grundsätzlich der Hierarchical Bayes-Ansatz der Choice-Based Conjoint-Analyse, wie von *Nitschke* und *Völckner* korrekt vorgeschlagen (S. 755f.).

Zur Vermeidung des ersten Unsicherheitsbias setzt man nun bei allen Stimuli des Befragungsdesigns die vom Probanden erwarteten Ausprägungen statt der in der Befragung extern vorgegebenen ein. Dazu muss man zunächst die Verteilung oder direkt die erwarteten Ausprägungen erfragen, wie ohnehin von *Nitschke* und *Völckner* vorgeschlagen (S. 749f.). Dadurch bekommt man für diese Person (bezüglich des ersten Bias) korrekte Teilnutzenwerte für die untersuchten (individuell erwarteten) und nicht für die falschen (vorgegebenen) Ausprägungsstufen. Der Unterschied besteht hierbei also lediglich darin, dass die Schätzung nicht anhand einer für alle Probanden einheitlichen und damit prinzipiell falsch gewählten Durchschnittsausprägung, sondern für jeden Probanden individuell erfolgt. Eine Aggregation kann erst danach erfolgen, wobei jedoch zu berücksichtigen ist, dass für jeden Probanden ein unterschiedliches Untersuchungsdesign notwendig geworden sein kann. Dies wiederum erschwert eine Aggregation der

Teilnutzenwerte über die Probanden hinweg, da die Ermittlung der individuellen Teilnutzenwerte aufgrund unterschiedlicher Bewertungsbasen der Teilnehmer erfolgte.⁶ Diesem Problem kann zum Beispiel dadurch begegnet werden, dass schon vorher Gruppen mit identischer merkmalsbezogener Basis gebildet werden, d.h. in diesem Sinne eine a priori Segmentierung vorgenommen wird.

Ist auch der zweite Bias von Bedeutung, so kann durchaus der Vorschlag von *Nitschke* und *Völckner* aufgegriffen werden, nicht die erwartete Ausprägung einzusetzen, sondern das Sicherheitsäquivalent der Ausprägungslotterie. Dabei sollte man allerdings darauf verzichten, (nicht begründete) Annahmen über den Verlauf der Risikonutzenfunktion zu treffen, wie auf Seite 752 beschrieben, sondern direkt das Sicherheitsäquivalent bezüglich der konkret zu beurteilenden Entscheidungsalternativen erheben. Dies ließe sich einfach zum Beispiel wie folgt bewerkstelligen: Man legt dem Entscheider zehn Filmangebote mit fein abgestuften *sicheren* Qualitäten von hoch bis niedrig über mehrere Zwischenstufen vor und fragt ihn, bei welchem Angebot er indifferent zu einer bestimmten Tauschbörse mit unsicherer Qualität ist (unter ceteris-paribus-Bedingungen). Damit erhält man ohne fragwürdige Annahmen und komplexe Berechnungen eine für die Nutzenschätzung verwendbare Ausprägung, die nutzenmäßig derjenigen entspricht, mit der er aufgrund seiner Risikoeinstellung rechnet.

Die vorhergehenden Ausführungen zeigen, dass eine explizite Berücksichtigung der Unsicherheit bereits während der Conjoint-Analyse prinzipiell möglich ist, aber diese verschiedene Schwierigkeiten verursacht. Besser ist es daher in der Regel, die Conjoint-Analyse zunächst unter vollkommener Sicherheit durchzuführen und die Unsicherheit nur bei der Berechnung der individuellen Gesamtnutzen für konkrete, mit Unsicherheit behaftete Angebote in der am Ende von Abschnitt 3 geschilderten korrekten Weise einzubeziehen. Dann muss gegebenenfalls auf herkömmliche Arten der Aggregation verzichtet werden, die jedoch in vielen Fällen auch gar nicht erforderlich sind. Statt der Berechnung aggregierter Gesamtnutzenwerte für (Produkt-) Angebote kann zum Beispiel die Zählung der individuellen Fälle treten, bei denen das eine oder andere Angebot besser bewertet wurde.

5. Sinn und Notwendigkeit expliziter Unsicherheitsberücksichtigung

Nun stellt sich allerdings die Frage, ob es überhaupt sinnvoll und notwendig ist, die durch Unsicherheit verursachten Unter- oder Überbewertungen von Produktmerkmalen aus den Ergebnissen einer Conjoint-Analyse zu eliminieren, um korrekte Nutzeinschätzungen für *tatsächlich* vorliegende Ausprägungen von Stimuli oder konkreten Pro-

⁶ Vgl. *Voeth* (2000), S. 177.

duktangeboten über alle Probanden aggregiert zu erhalten. Dies ist sicherlich sinnvoll, wenn es um Produkte mit Sucheigenschaften geht, bei denen aber eine Ausprägungsunsicherheit definitionsgemäß praktisch nicht auftreten kann (S. 744).

Liegen jedoch Erfahrungs- und Vertrauenseigenschaften vor, so ist es notwendig, die Existenz von Unsicherheit und den individuellen Umgang damit im Rahmen der Präferenzmessung mit einzubeziehen, um vollständige Kaufentscheidungsprozesse abbilden zu können.⁷ Die Präferenzanalyse soll dabei Entscheidungen möglichst gut vorhersagen, d.h. Präferenzmessmethoden müssen die Entscheidungsprozesse bestmöglich abbilden. Extensive Kaufentscheidungsprozesse bzw. die letzten Phasen innerhalb des evoked sets und dementsprechend Präferenzmessmethoden basieren auf Trade-offs,⁸ d.h. die Nachfrager bewerten Eigenschaftsausprägungen verschiedener Alternativen des evoked sets⁹ – sie setzen sie zueinander in Beziehung. In diesen Trade-offs sind aber die Unsicherheiten bei der Bewertung inhärenter Bestandteil. Schließlich kann der Kunde die ideale Informationslage vor dem Kauf nicht herstellen, so dass seine Bewertung der tatsächlichen Ausprägungen für den realen Kaufprozess gar keine Rolle spielt, sondern nur seine Präferenzen für die *individuell vermuteten* Eigenschaften unter der individuellen Risikoeinschätzung. Insofern würde eine Conjoint-Analyse unter diesen Bedingungen eine Einschätzung der individuellen Präferenzen und individueller Kaufwahrscheinlichkeiten erlauben, jedoch wäre eine Aggregation der Ergebnisse nicht ohne weiteres möglich, da bei jedem Probanden individuelle Ausprägungsstufen zugrunde liegen. Um dies dennoch zu tun, müsste man die verwendeten Ausprägungsstufen individuell erfragen und voneinander bei der Teilnutzenschätzung differenzieren, wie in Abschnitt 4 vorgeschlagen.

In realen Situationen entscheiden Käufer regelmäßig, ob sie weitere Informationen brauchen und somit ihre Entscheidung besser fundieren müssen. Tun sie dies nicht, sind die Teilnutzenwerte *ceteris paribus* geringer, da ihnen das risikobehaftete Merkmal offensichtlich nicht wichtig genug ist. Holen sie dagegen weitere Informationen ein, halten sie es aufgrund der Bedeutung des Merkmals für wichtig, ihr Risiko zu reduzieren. Konsequenterweise sollte eine Auskunftsperson / ein Nachfrager vor der Präferenzbildung und -messung also entscheiden können, ob sie / er zusätzliche Informationen braucht (in der Messung beispielsweise mit einem information display board). Es sollte dann nicht im Nachhinein eine vermutete Unsicherheit in das Entscheidungskalkül integriert werden. Stattdessen sollte eine Unsicherheitsreduktion durch a priori Einteilung der Auskunftspersonen gemäß ihrer Präferenzstrukturen so vorgenommen werden, dass

⁷ Vgl. Backhaus/Brzoska (2004), S. 41.

⁸ Vgl. Green/Srinivasan (1978), S. 104.

⁹ Vgl. Mehta/Moore/Pavia (1992), S. 470.

jeder Gruppe nur individuell wichtige Merkmale und solche, die sie geeignet beurteilen können und wollen, vorgelegt werden.¹⁰

Es ist also durchaus fraglich, ob es in bestimmten Fällen – im Sinne der realistischen Abbildung realer Entscheidungssituationen nicht sogar günstiger ist – die Unsicherheit im Rahmen der Conjoint-Analyse zuzulassen, um genau deren Auswirkungen auf die Bewertungen und die dadurch simulierte Kaufentscheidung realitätsnah abzubilden.

6. Zur empirischen Studie

Nitschke und *Völckner* beschreiben eine Studie (S. 754ff.), die aufzeigen soll, dass Unsicherheit tatsächlich zu Nutzenwertabweichungen führt und ob Kunden bereit sind, für sichere Angebote mehr zu bezahlen. Neben den bereits diskutierten methodischen Mängeln des gesamten Ansatzes weist die Studie weitere Mängel auf, die das Erreichen dieser Ziele zusätzlich in Frage stellen. Die wichtigsten seien kurz erwähnt:

Im Rahmen der Präferenzmessung auf Basis additiver Gesamtnutzenfunktionen müssen die Eigenschaften voneinander unabhängig sein,¹¹ d.h. es dürfen keine schätzparameterverzerrenden Interaktionseffekte auftreten.¹² Jedoch sind gerade die in der Studie zentralen Eigenschaften Preis und Qualität in der Literatur das Paradebeispiel dafür, was hinsichtlich möglicher Korrelationen von Merkmalen unbedingt vermieden werden müsste. Hier gibt es gängige Lösungsmöglichkeiten, wie beispielsweise die Integration des Preises als abhängige Variable oder die Nutzung von Compound-Attributes. Aufgrund der Nichtbeachtung dieser Interaktionen sind unter anderem auch vor diesem Hintergrund die empirischen Befunde fragwürdig.

Die Vorgehensweise bei der Berechnung mittlerer Ausprägungen anhand der erhobenen Wahrscheinlichkeitsverteilung (S. 757f.) ist nicht angemessen, da ordinal skalierte Ausprägungen zur Erwartungswertberechnung als intervallskaliert interpretiert werden. Eine derartige Anhebung des Skalenniveaus ist messtheoretisch falsch, auch wenn eine solche Vorgehensweise auch bei anderen Studien verwendet wurde (S. 758). Möchte man aber gerade Verzerrungen durch Messung falscher Teilnutzenwerte vermeiden, ist ein solches Vorgehen nicht anzuraten. Stattdessen sollte man von vornherein eine feinere Abstufung der Ausprägungsstufen vorsehen.

Von der Messung einer allgemeinen Risikoeinstellung, wie auf Seite 766f. gezeigt, kann nicht auf eine produktbereichsspezifische Risikoeinstellung geschlossen werden – diese

¹⁰ Vgl. *Hauser/Urban* (1977), S. 589; *Ratneshwar/Shocker/Stewart* (1987), S. 523f.

¹¹ Vgl. *Green/Helsen/Shandler* (1988), S. 392.

¹² Vgl. *Leigh/MacKay/Summers* (1984), S. 456.

kann von vorhandenen Erfahrungen bzw. vom Involvement im Produktbereich abhängen und damit von der allgemeinen Risikoeinstellung erheblich abweichen. Außerdem weisen die betrachteten monetären Werte im Experiment zur Risikoeinstellung einerseits und innerhalb des eigentlichen Problems andererseits unterschiedliche Größenordnungen (5 bis 50 Euro versus 0 bis 3,60 Euro) auf, so dass eine Übertragung nicht statthaft ist. Ein Entscheider kann nämlich zum Beispiel bei 3 Euro risikofreudig und bei 50 Euro risikoscheu agieren. Hinzu kommen unnötige fragwürdige Annahmen über Typ und Verlaufsform der Risikonutzenfunktion (S. 752).

7. Fazit

Abschließend sei gesagt, dass ein Verdienst des Artikels von *Nitschke* und *Völckner* sicherlich darin besteht, auf die mit Ausprägungsunsicherheit einhergehenden Schwierigkeiten hinzuweisen. Diese Problematik wurde bislang absolut unzureichend behandelt. Die Grundlinien des Vorschlags, eine Verteilung der unsicheren Ausprägungsstufen zu erheben und die Unsicherheit mit entscheidungstheoretischen Konzepten in die Präferenzmessung zu integrieren, sind zu begrüßen. Weitere Forschung sollte darauf aufbauen.

Jedoch kann nicht übersehen werden, dass der Artikel an zentralen Stellen durch falsche Ausgestaltung der Conjoint-Analyse und/oder falsche Schlussfolgerungen zu erheblicher Verwirrung beiträgt. Diese wird (hoffentlich) durch die vorgelegten Anmerkungen deutlich reduziert.

Literatur

Backhaus, K.; Brzoska, L. (2004): Conjointanalytische Präferenzmessung zur Prognose von Preisreaktionen, in: *Die Betriebswirtschaft (DBW)*, Vol. 64, Nr. 1, S. 39-57.

Backhaus, K.; Erichson, B.; Plinke, W.; Weiber, R. (2003): *Multivariate Analysemethoden*, 10. Aufl., Berlin.

Gierl, H. (2006): Präferenzen bei Ergebnisambiguität, in: *Zeitschrift für Betriebswirtschaft (ZfB)*, Vol. 76, No. 1, S. 1187-1216.

Green, P.E.; Helsen, K.; Shandler, B. (1988): Conjoint Internal Validity under Alternative Profile Presentations, in: *Journal of Consumer Research*, Vol. 15, No. 3, S. 392-397.

Green, P.E.; Srinivasan, V.S. (1978): Conjoint Analysis in Consumer Research: Issues and Outlook, in: *Journal of Consumer Research*, Vol. 5, No. 2, S. 103-123.

Hauser, J.R.; Urban, G.L. (1977): A Normative Methodology for Modeling Consumer Response to Innovation, in: *Operations Research*, Vol. 25, No. 4, S. 579-619.

- Keeney, R.L.; Gregory, R.S.* (2005): Selecting Attributes to Measure the Achievement of Objectives, in: *Operations Research*, Vol. 53, No. 1, S. 1-11.
- Klein, R.; Scholl, A.* (2004): *Planung und Entscheidung*, München.
- Leigh, T.W.; MacKay, D.B.; Summers, J.O.* (1984): Reliability and Validity of Conjoint-Analysis and Self-Explicated Weights: A Comparison, in: *Journal of Marketing Research*, Vol. 21, No. 4, S. 452-462.
- Mehta, R.; Moore, W.L.; Pavia, T.M.* (1992): An Examination of the Use of Unacceptable Levels in Conjoint Analysis, in: *Journal of Consumer Research*, Vol. 19, No. 3, S. 470-476.
- Nitschke, T.; Völckner, F.* (2006): Präferenzmessung bei unsicheren Produkteigenschaften: Risikoberücksichtigung bei Ergebnissen aus Conjoint-Analysen, in: *zfbf*, 58. Jg., S. 743-770.
- Payne, J.W.; Bettman, J.R.; Schkade, D.A.* (1999): Measuring Constructed Preferences: Toward a Building Code, in: *Journal of Risk and Uncertainty*, Vol. 19, No. 1, S. 243-270.
- Ratneshwar, S.; Shocker, A.D.; Stewart, D.W.* (1987): Toward Understanding the Attraction Effect: The Implications of Product Stimulus Meaningfulness and Familiarity, in: *Journal of Consumer Research*, Vol. 13, No. 4, S. 520-533.
- Shocker, A.D.; Srinivasan, V.S.* (1979): Multiattribute Approaches for Product Concept Evaluation and Generation: A Critical Review, in: *Shocker, A.D.* (Hrsg.), *Analytic Approaches to Product and Marketing Planning*, Cambridge, S. 3-38.
- Voeth, M.* (2000): *Nutzenmessung in der Kaufverhaltensforschung: Die Hierarchische Individualisierte Limit Conjoint-Analyse (HILCA)*, Wiesbaden.

Zusammenfassung

Der Beitrag beschäftigt sich mit der Problematik, dass im Rahmen von Präferenzmessungen Ausprägungen einzelner Merkmale unsicher sein können. Die möglichen negativen Auswirkungen dieser Unsicherheit und Ansätze zu ihrer Beseitigung im Rahmen von Conjoint-Analysen werden in einem Beitrag von Nitschke und Völckner thematisiert. Die dort gemachten Vorschläge und Schlussfolgerungen werden hier einer kritischen Prüfung unterzogen. Dabei stellt sich heraus, dass die beschriebenen Effekte zwar bedeutsam sind, die vorgeschlagene Methodik jedoch nicht zu deren Berücksichtigung geeignet ist. Korrekturen und Alternativvorschläge werden unterbreitet.

Summary

Within preference measurement, uncertainty of attribute levels describing alternatives to be evaluated is a common problem which might cause deviations between measured utility functions and actual ones. In a paper by Nitschke and Völckner, published in *Zeitschrift für betriebswirtschaftliche Forschung* (Vol. 58/9, 743-770), negative consequences of uncertain attribute levels and solution approaches are discussed. While the effects shown are correct and important, the proposals to overcome them by correcting utility values in an ex post manner turn out to be insufficient and inappropriate. These findings are discussed in the note on hand which also contains corrections and alternative solution approaches.